

The **Anatomy of a Decision:** Inside the SahAI Agentic Core

An Architectural Deep Dive into Voice-First,
Hindi-Native AI Reasoning

The Journey Begins with a Single Voice

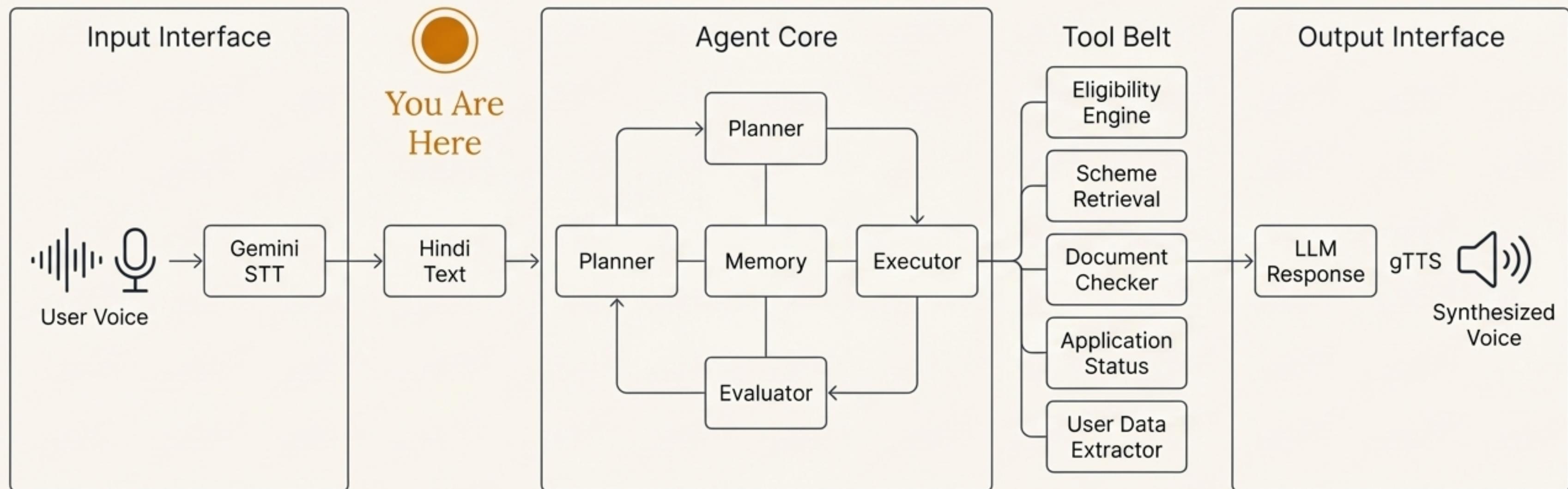


मैं एक 65 वर्षीय किसान हूँ, मेरी आय
50,000 रुपये है। क्या मैं किसी योजना के
लिए पात्र हूँ?

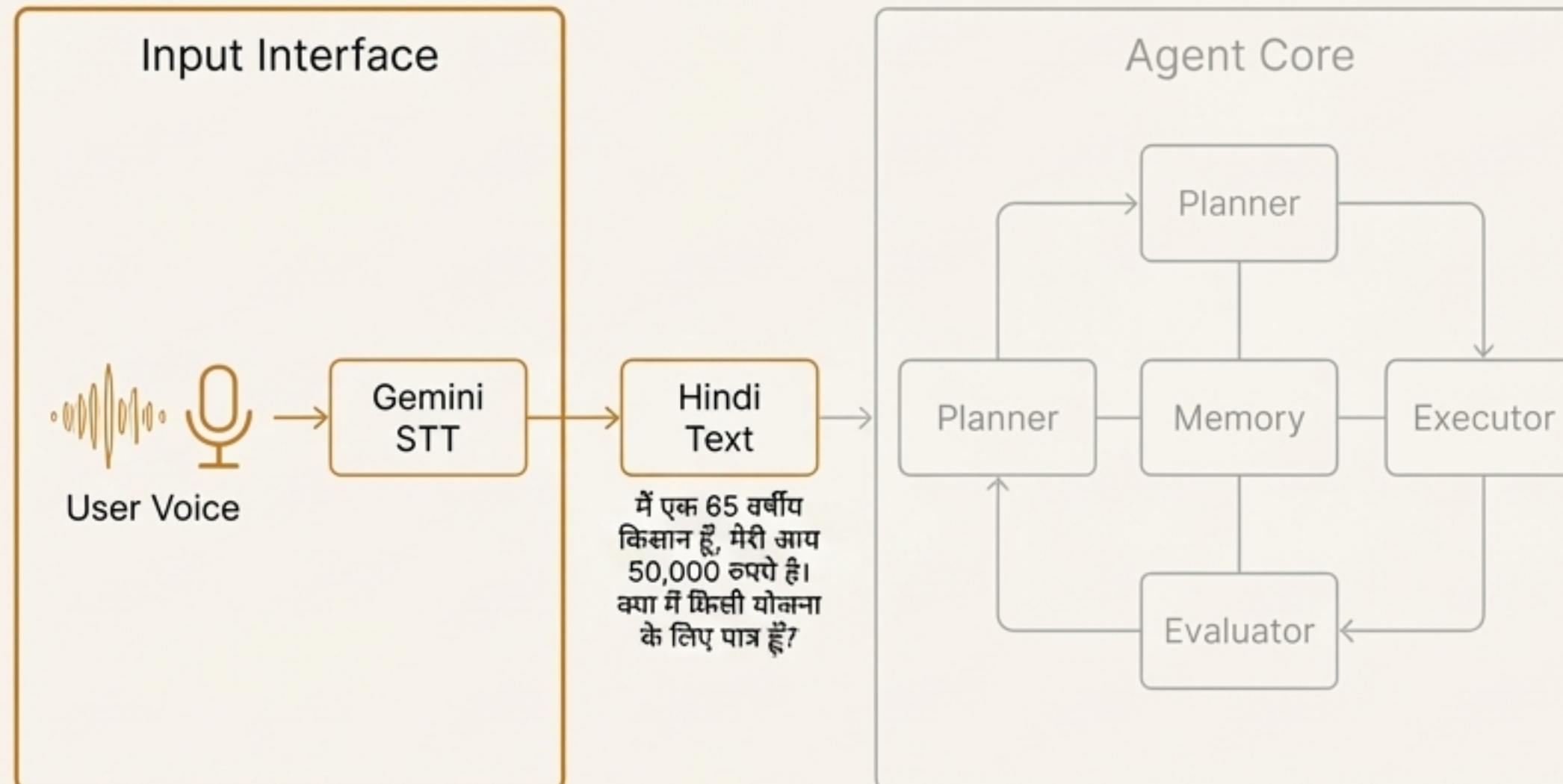
(I am a 65-year-old farmer with an income of
50,000 rupees. Am I eligible for any schemes?)

This isn't a simple query. It's a multi-faceted request requiring data extraction, rule-based checks, and information retrieval. Our architecture is designed to deconstruct and solve it.

The SahAI Cognitive Architecture: A High-Level Blueprint

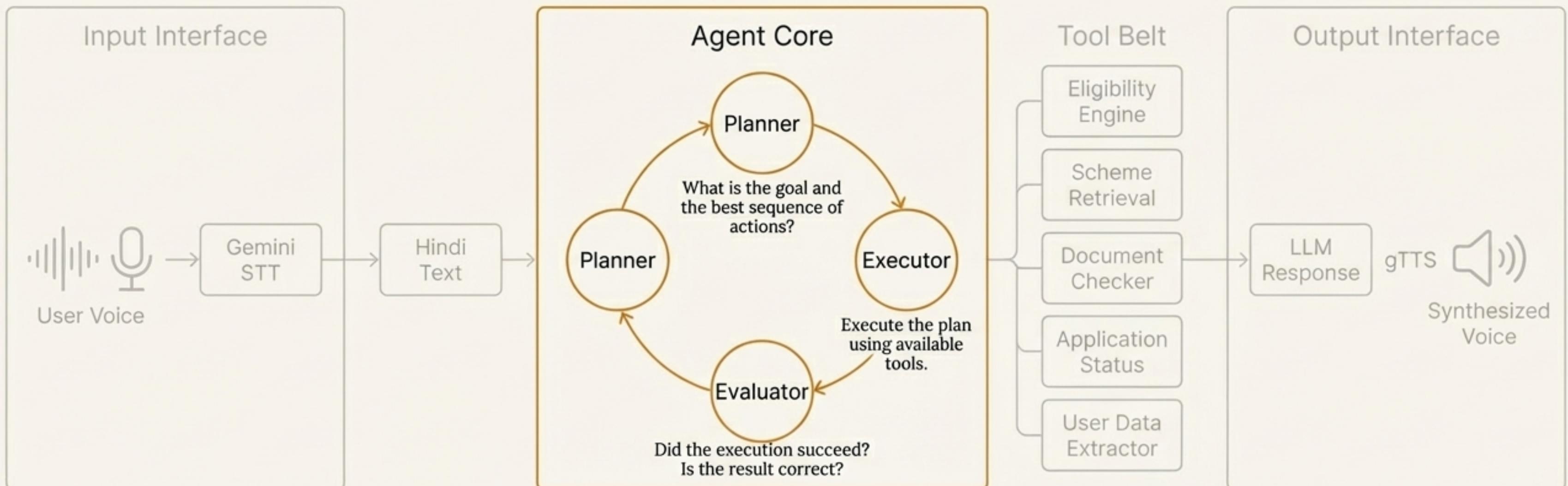


Step 1: From Sound to Signal – Hindi Speech-to-Text



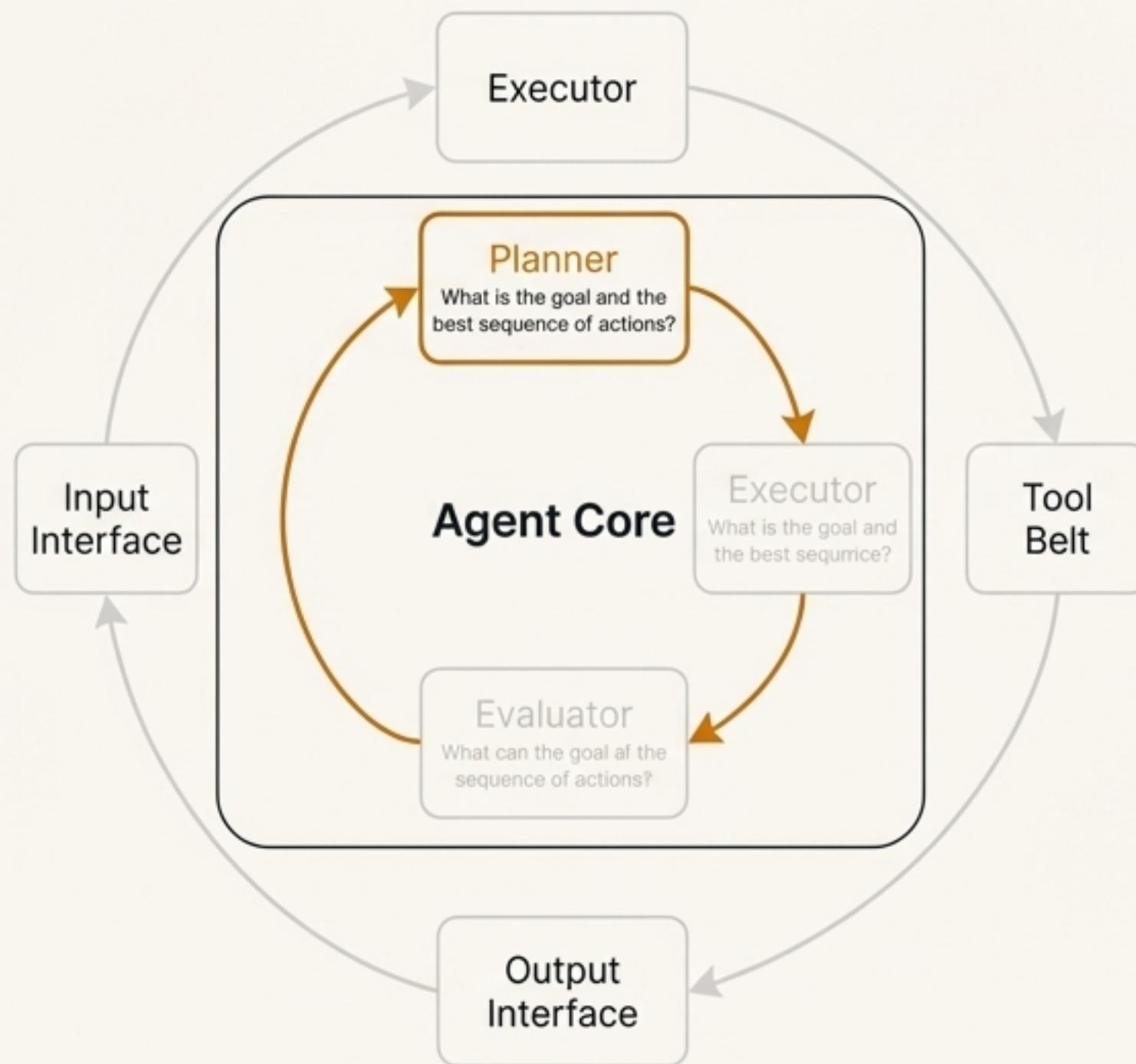
- * Primary Modality: Designed for voice-first interaction.
- * Technology: Leverages Gemini STT for high-accuracy Hindi transcription.
- * Resilience: Includes robust error handling for no audio detection, unclear speech, and partial transcriptions.

The Agent's Core: A Deliberate Reasoning Loop



This Planner-Executor-Evaluator loop enables the agent to reason, self-correct, and handle complex, multi-step tasks dynamically.

The Planner: Deconstructing Intent and Formulating a Strategy



PLANNER PROMPT (Conceptual)

CONTEXT

User Query: "मैं एक 65 वर्षीय किसान हूँ, मेरी आय 50,000 रुपये है। क्या मैं किसी योजना के लिए पात्र हूँ?"

Available Tools: [Eligibility Engine, Scheme Retrieval, Document Checker, ...]

TASK

1. Identify the user's ultimate goal.
2. Extract all relevant entities (age, income, profession).
3. Formulate a step-by-step plan using the available tools to achieve the user's goal.

OUTPUT (PLAN)

Key Insight: The Planner's job is not to answer the question, but to create a precise, executable plan for how to answer it.

The Executor: Translating the Plan into Action



Tool Belt



Eligibility Engine

Checks user eligibility against scheme criteria.



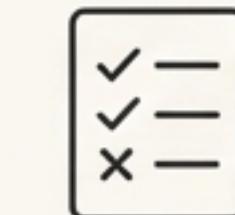
User Data Extractor

Extracts structured data (age, income, gender) from Hindi text.



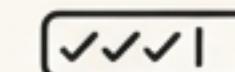
Scheme Retrieval

Searches and retrieves scheme information.



Document Checker

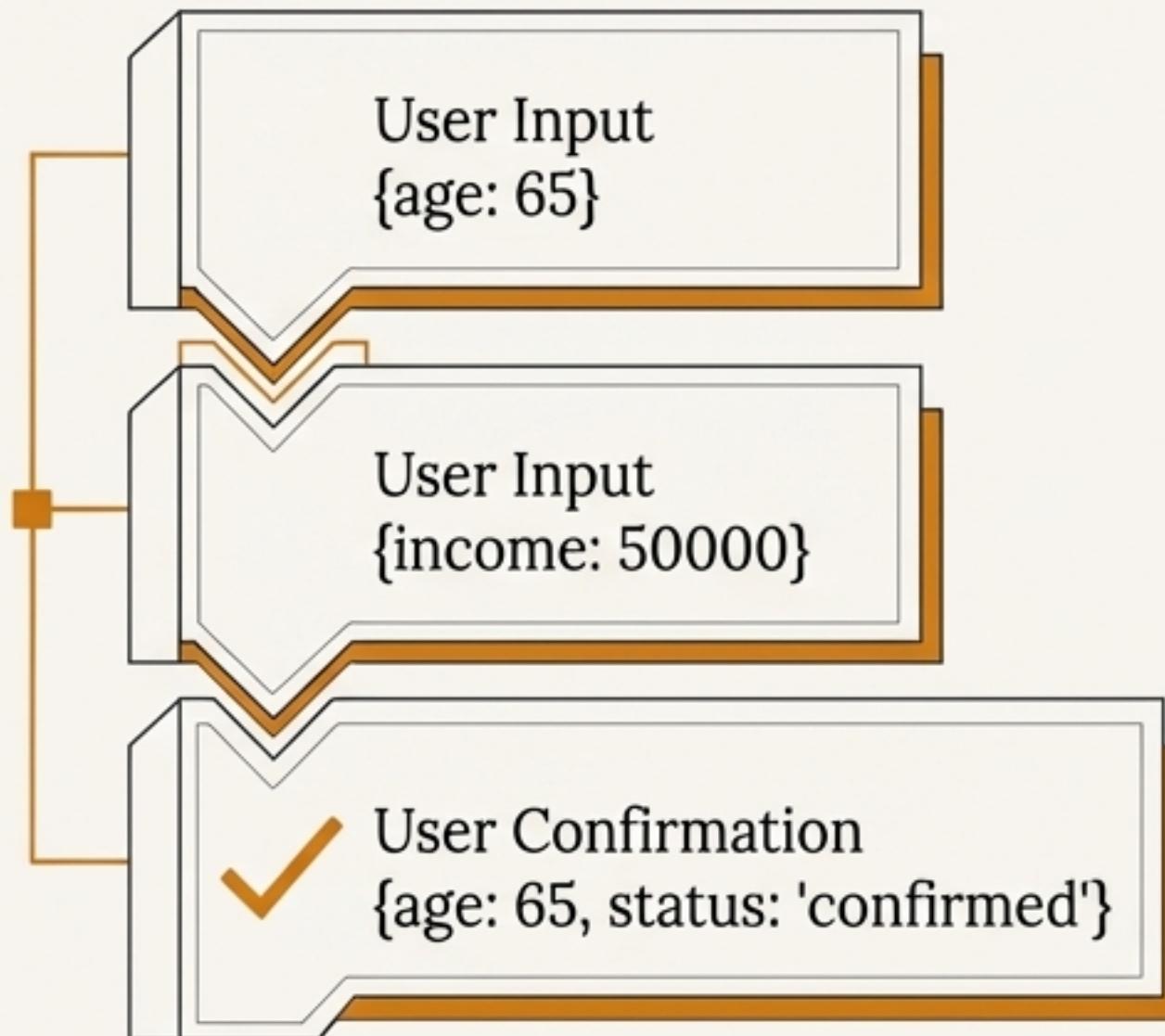
Lists required documents for each scheme.



Application Status

Simulates checking application status (Mock API).

The Agent's Memory: Maintaining Context and Ensuring Coherence



- **Session-based Tracking:** Remembers user data (age: 65, income: 50000) throughout the conversation.



- **Field History:** Knows when a piece of information was provided.



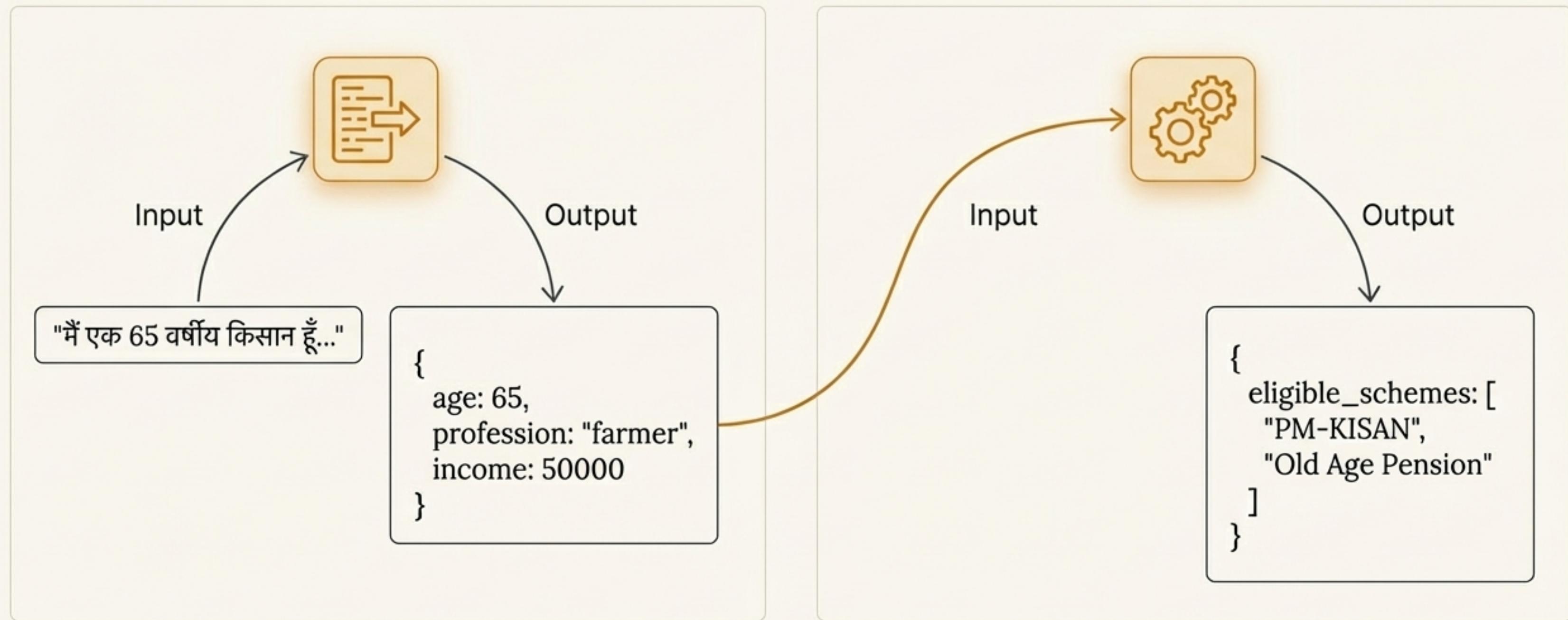
- **Contradiction Detection:** Flags conflicting information (e.g., if the user later says their age is 45).



- **Confirmation Tracking:** Marks data points as confirmed by the user.

This is not just a simple chat history; it's a **structured, stateful memory** that underpins the agent's reasoning and reliability.

Execution in Progress: Applying Tools to Data



The Executor successfully used two tools to transform an unstructured query into a structured, actionable result.

The Evaluator: A Final Check for Accuracy and Success



EVALUATOR PROMPT (Conceptual)

CONTEXT

Original Goal: Check eligibility for a 65-year-old farmer with 50k income.

Plan: 1. Extract data. 2. Run Eligibility Engine.

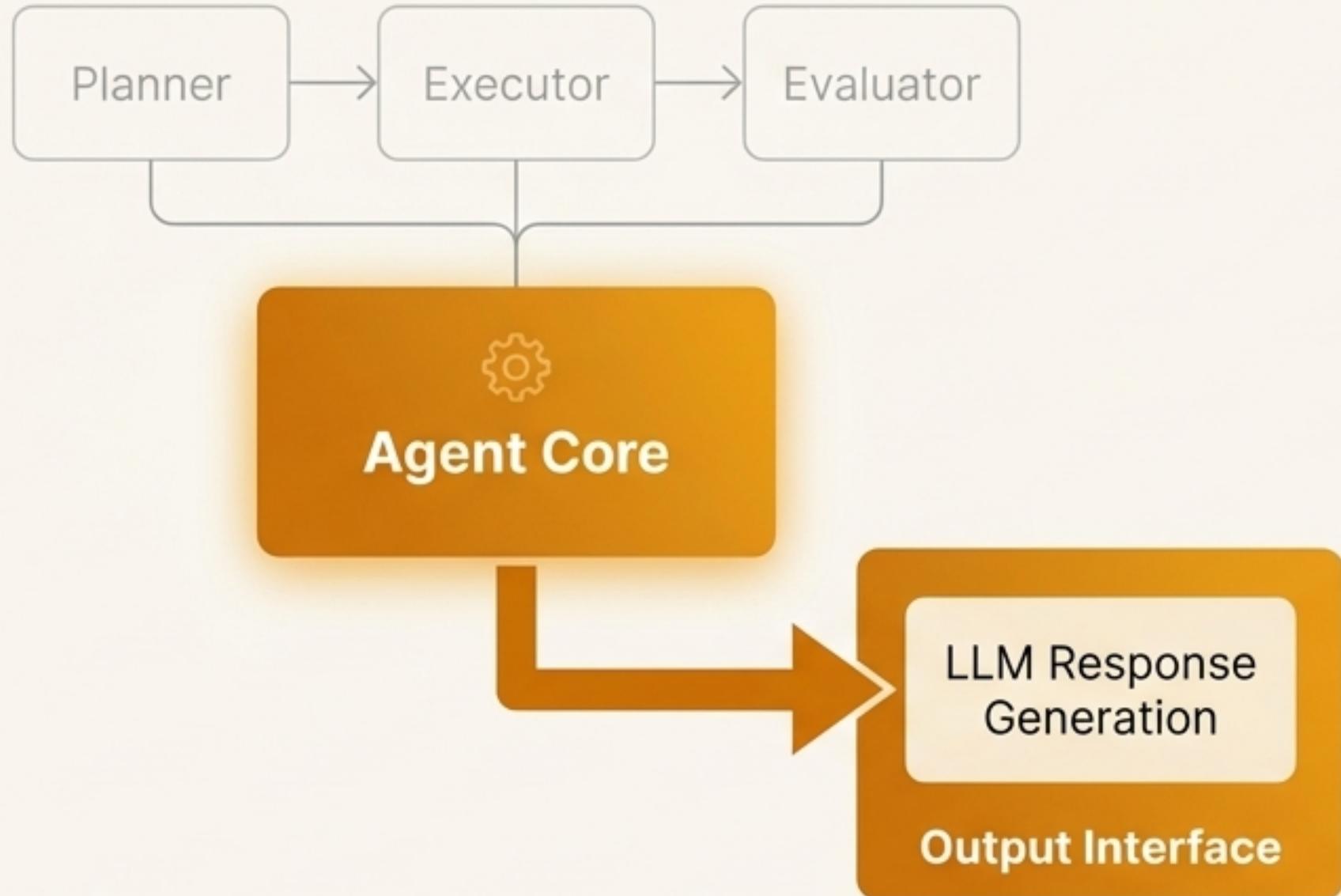
Execution Result: Found 2 eligible schemes: PM-KISAN, Old Age Pension.

TASK

1. Does the execution result directly satisfy the original goal? (YES/NO)
2. Is the information complete and sufficient to answer the user? (YES/NO)
3. If NO, identify the failure and suggest a revised plan.

Key Insight: The Evaluator closes the loop. It ensures the final output is relevant and correct, enabling the agent to re-plan if necessary.

Step 2: Synthesizing a Coherent Response

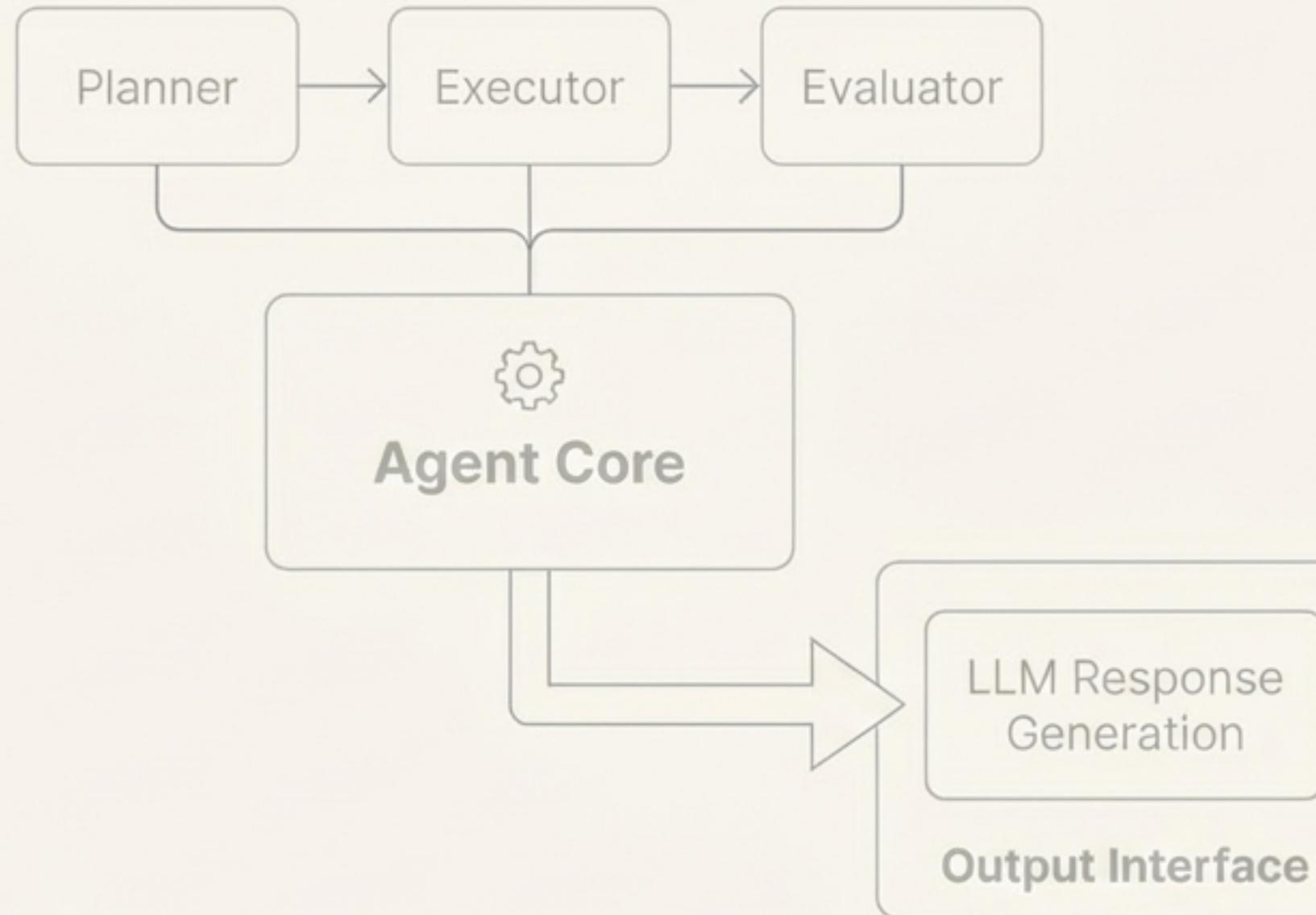


Generated Text Response

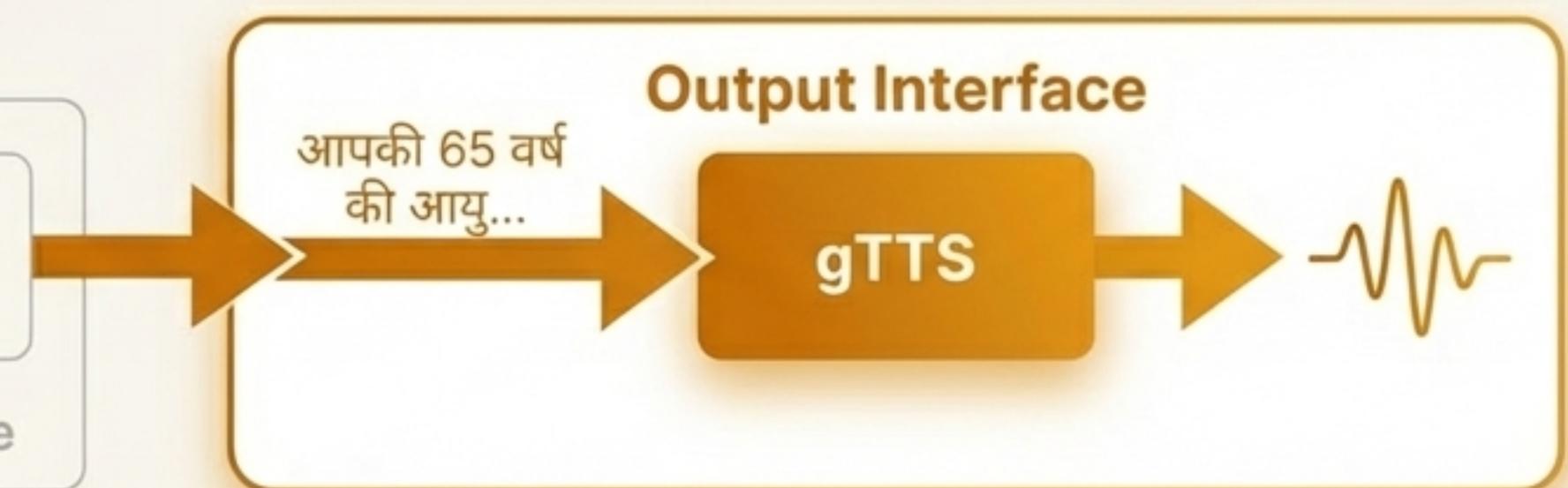
आपकी 65 वर्ष की आयु और 50,000 रुपये की आय के आधार पर, आप दो योजनाओं के लिए पात्र हो सकते हैं: पीएम किसान सम्मान निधि और सम्मान निधि, और वृद्धावस्था पेंशन योजना। क्या आप किसी एक के बारे में और जानना चाहेंगे?

(Based on your age of 65 and income of 50,000 rupees, you may be eligible for two schemes: PM-KISAN and the Old Age Pension scheme. Would you like to know more about either one?)

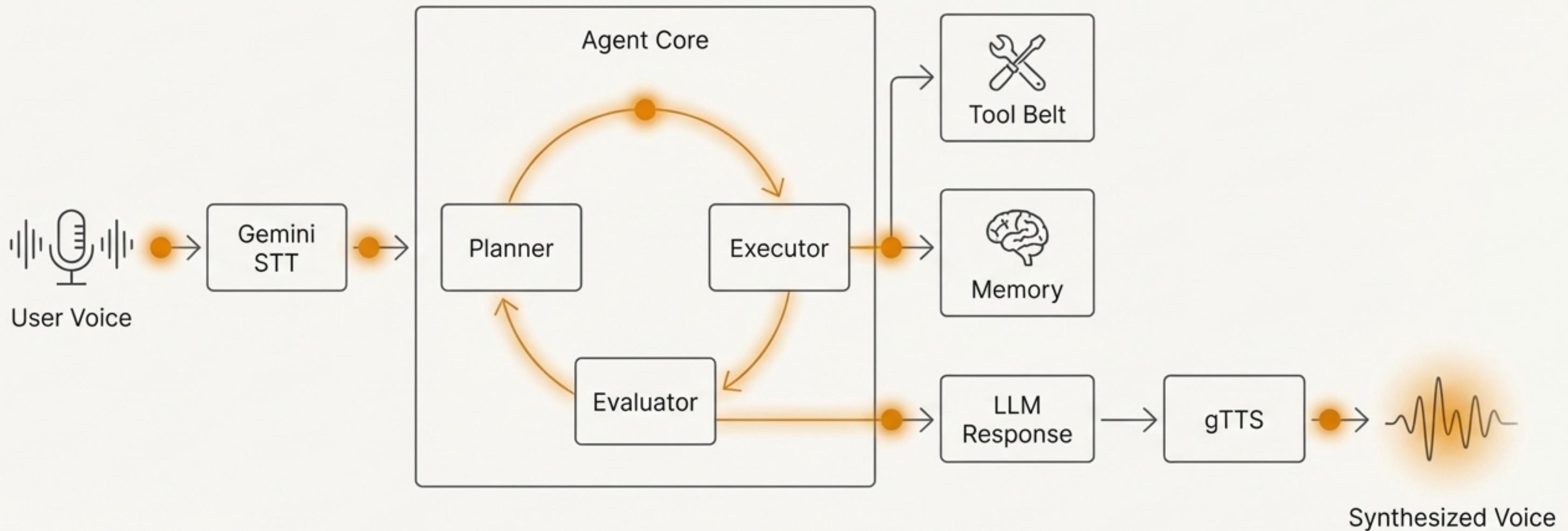
Step 3: From Signal to Sound – Hindi Text-to-Speech



- **Technology:** Utilizes gTTS for clear and natural-sounding Hindi speech.
- **Seamless Pipeline:** Completes the full STT → LLM → TTS cycle in native Hindi.
- **Accessibility:** Ensures the system is truly voice-first and accessible to a wide range of users.



The Anatomy of a Decision: The Full Journey



Architected for Intelligence and Trust



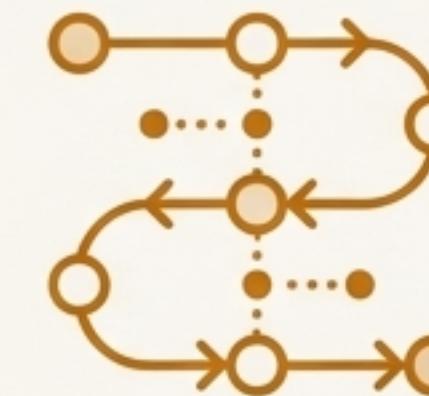
True Agentic Reasoning

The Planner-Executor-Evaluator loop enables dynamic problem-solving and self-correction, moving beyond simple Q&A.



Extensible Tool Use

A modular set of 5 distinct tools allows the agent to interact with systems and perform complex, real-world tasks.



Stateful Conversational Memory

Contradiction detection and history tracking ensure coherent, context-aware interactions over multiple turns.



Comprehensive Failure Handling

Graceful recovery from STT errors, missing information, and system issues builds user trust and system reliability.